DRAFT

SECTION J - BENCHMARK INSTRUCTIONS

The offeror must complete the following tables related to the performance of the benchmark suite on their offered production and backup systems.

The capability benchmark codes consist of the global forecast model GFS T382, the global analysis GSI T382, the mesoscale forecast model NMM 8km. The capacity benchmark is the GFS T126. The offeror must provide measurements for three distinct configurations to demonstrate scalability of the capability benchmark codes on a single system in Tables 1a and 1b. The offeror must describe the configuration of each system, with respect to architecture, number of processors and amount of memory. The capability and capacity benchmarks will make up a major portion of the LTD suite.

Table 1a should contain wall clock run times in seconds for the capability benchmark codes. This table is intended to aid the government in determining the risk associated with performance projections throughout the life of the contract.

Table 1b should contain the number of forecast days completed in a 60-minute window for the capacity benchmarks GFS T126 with high IO and GFS T126 with low IO requirements. High IO is defined as an output frequency of every hour and low IO is defined as an output frequency of every 24 hours. Although the GFS T126 is a capacity benchmark, the government is interested in the scalability of the GFS T126 on the proposed system under various system configurations.

Table 2a should contain either actual or projected system performance as a speedup factor of the preceding operational system running two concurrent copies. The IBM Baseline run times contained in column 2 were measured on the current IBM production system.

Table 2b should contain either actual or projected system performance as an increase factor compared to the preceding operational system Forecast days in a 60 minute window. The IBM Baseline Forecast Days was measured on the current IBM production system.   The same executable must be used with the IO rate controlled at the script level. Each single execution of the capacity benchmark will run for 3 days whereupon another copy can be initiated from the queue. The jobs must be initiated by the vendor's proposed batch system. The vendor is encouraged to run as many concurrent copies as possible to maximize the throughput and increase the total number of forecast days. The initial delivery column is the speedup factor over the IBM Baseline and the phase II column is the speedup factor over the initial delivery column.

Table 3 should contain the projected percentage of the offered system that would be required to run the capability benchmark within the wall clock time window provided. Table 3 will aid the government to evaluate the capability of the projected system in meeting operational schedules.

Tables 4 and 5 should contain the projected performance guarantees for each phase with respect to the preceding phase. The minimum capability performance guarantee must be met by each of the capability benchmarks. A speedup factor must be computed for each capability benchmark by running 2 concurrent copies. The average capability performance guarantee must be met by averaging the three capability benchmarks. The capacity performance guarantee for the high IO must be met by the capacity benchmark using the high IO script. The capacity performance guarantee for the low IO must be met by the capacity benchmark using the low IO script. The low and high IO capacity scripts use the same executable.

The offeror will be allowed to make software improvements, subject to the requirements in Section E for the initial delivery only. The government will make a risk assessment of source code changes. The government strongly discourages any source code changes that reduce portability. The baseline numbers for the initial delivery will be the IBM numbers provided in Table 1a and Table 1b.

Future upgrades beyond the initial delivery (10/05) will be based on conditions existing one month prior to scheduled delivery.
1. The capability benchmark suite will consist of the operational versions of the global forecast, global analysis, meso forecast, meso analysis, and ocean model one month prior to scheduled delivery.
2. The capacity benchmark will be the operational version of the global forecast model at approximately one quarter of the operational resolution one month prior to scheduled delivery.
3. A new baseline will be generated on the operational system for each benchmark to use for measuring the next speedup guarantee.
4. No code changes or script changes will be allowed without government concurrence.

Table 1a : Capability Benchmark Performance under various system configurations.

Wall clock seconds	IBM Baseline nodes/cpus/time	Configuration 1	Configuration 2	Configuration 3
Global Fcst GFS T382				
Global Analysis GSI T382				
Meso Fcst NMM 8km				

Table 1b : Capacity Benchmark Performance under various system configurations

Forecast Days	IBM Baseline	Configuration 1	Configuration 2	Configuration 3
GFS T126 high IO				
GFS T126 low IO				

Table 2a: Projected Capability system performance with respect to the preceding system.

X times IBM Baseline	IBM Baseline (Seconds)	Initial (Oct 2005) secs/X-factor	Phase II secs/X-factor
Global Fcst GFS T382			
Global Analysis GSI T382			
Meso Fcst NMM 8km			
Average			

Table 2b: Projected Capacity system performance with respect to the preceding system.

X times IBM Baseline	IBM Baseline (Fcst Days in 30 minute window)	Initial (Oct 2005) fcst days/X-factor	Phase II fcst days/X-factor
GFS T126 high IO			
GFS T126 low IO			

Table 3: Percentage of the initial system required to satisfied wall clock requirement for benchmark runs

CPU Resource	Time Window	IBM Operational %	Initial Delivery %
Global Fcst GFS T382(24hr)	10 minutes		
Global Analysis GSI T382	20 minutes		
Meso Fcst WRF 8km (24 hr)	24 minutes		

Table 4: Base Performance Guarantee (X times preceding phase operational baseline)

Phase	Date	Minimum Capability	Average Capability	Capacity high IO	Capacity low IO
Base Initial	04/01/06				
Phase II	04/01/08				

Table 5: Option Performance Guarantee (X times preceding phase operational baseline)

Phase	Date	Minimum Capability	Average Capability	Capacity high IO	Capacity low IO
Option Initial	10/01/10				
Phase II	10/01/12				